



UNIVERSIDAD
COMPLUTENSE
MADRID

Proyecto de Innovación

Convocatoria: 2017/2018

Nº de proyecto: 082

Título de proyecto:

“Recursos interactivos para prácticas con Software Estadístico en Informática”

Nombre del responsable del proyecto: Alba María Franco Pereira

Centro: Facultad de Ciencias Matemáticas

Departamento: Estadística e Investigación Operativa

Índice

Objetivos propuestos en la presentación del proyecto.....	3
Objetivos alcanzados.....	4
Metodología empleada en el proyecto.....	6
Recursos humanos.....	7
Desarrollo de las actividades.....	8
Anexo.....	9

1. Objetivos propuestos en la presentación del proyecto

La asignatura de Estadística de nuestros ingenieros informáticos presenta un alto número de suspensos así como de alumnos No Presentados. En los últimos años, el porcentaje de Suspensos en relación al número de Matriculados ha sido del 37%, el de Suspensos frente a Presentados alrededor del 50% y el de No Presentados de los Matriculados un 25%. Por otra parte, la Tasa de Rendimiento en los últimos años está alrededor del 45% y la Tasa de Éxito es un poco superior al 50%.

Los estudiantes de las titulaciones de Informática utilizan internet y distintos softwares en aquellas asignaturas de ámbito computacional. El principal objetivo del proyecto era mostrar a los alumnos de las asignaturas de Estadística la utilidad de las técnicas que estudian durante el curso. En este proyecto nos hemos centrado en la aplicación de la Estadística Descriptiva y el Análisis de Regresión.

En particular, los objetivos que nos trazamos en un inicio fueron los siguientes:

- Elegir entre los recursos informáticos disponibles en la UCM el más adecuado para estos alumnos en función de la dificultad y la aplicabilidad en estos fines en concreto.
- Elaborar un manual para el análisis estadístico (descriptivo y de regresión) a través de dicho software.
- Crear un banco de ejercicios y casos prácticos para utilizar en años posteriores.
- Incorporar a la virtualización de la asignatura los casos prácticos que permitan la evaluación de los distintos bloques temáticos.

Nuestra idea era clara, mejorar el grado de implicación del estudiante, modernizar el material didáctico e incentivar la investigación, el interés y el espíritu crítico de nuestros estudiantes. Para ello creemos necesario dar más visibilidad a la parte aplicada de la asignatura y que el alumno fuese capaz de, no sólo resolver problemas, sino también de plantearlos.

2. Objetivos alcanzados

El grupo de profesores que formamos el proyecto nos reunimos en varias ocasiones para coordinar el trabajo a realizar. En primer lugar nos planteamos diversos programas que podríamos utilizar para llevar a cabo las prácticas. Finalmente nos decidimos por Excel por su fácil manejo, versatilidad y familiaridad. Aunque no lo sabíamos en un principio, nos sorprendió que un porcentaje significativo de alumnos no había usado nunca Excel.

Como ya hemos comentado, los estudiantes de las titulaciones de Informática utilizan internet y distintos softwares en aquellas asignaturas de ámbito computacional. Con el propósito de fomentar este uso y darle el máximo valor añadido posible, queremos enfocar la asignatura de Estadística de manera que a los estudiantes les resulte atractiva y vean la utilidad de la estadística como herramienta en su contexto. Por este motivo creemos en la incorporación de software matemático y estadístico como parte integral de los currículos formativos.

Como se puede ver en el apartado **6. Anexos**, hemos elaborado un manual centrándonos en los temas que comentábamos en un inicio, y tal y como planteamos en los objetivos del proyecto: Estadística Descriptiva y Análisis de Regresión.

Respecto a la idea de crear un banco de ejercicios y casos prácticos aún está en elaboración. Nuestra idea es considerar los ejercicios entregados por los alumnos (entraremos más en detalle en el apartado **3. Metodología empleada en el proyecto**) que nos parezcan más relevantes incluirlos como parte del manual. Esta tarea la estamos realizando ahora mismo, pero nos gustaría seguir implementando la práctica tal y como hemos hecho en este curso y, una vez que tengamos un volumen mayor de ejercicios, crear el banco de ejercicios y ponerlo a disposición de los alumnos a través del campus virtual.

Observando el seguimiento de la actividad docente de estos alumnos de los grados de informática a lo largo de los últimos 6 años, debemos destacar cómo la incorporación de material docente en el campus virtual, la incorporación de algunos test interactivos, etc. ha mejorado la tasa de Presentados y de notas medias de los alumnos. No obstante, no es suficiente. Creemos que la elaboración de las Guías Docentes y de un sistema de evaluación práctica de la asignatura puede mejorar notablemente la tasa de Presentados y la tasa de notas medias altas en la asignatura. Por eso hemos llevado a cabo este proyecto.

La línea de trabajo basada en la elaboración de una guía docente o cuadernillo de prácticas no es del todo innovadora ya que en otras asignaturas de Estadística en otros grados también se ha llevado a cabo este enfoque. Concretamente en la asignatura de Estadística del grado de Farmacia, los alumnos utilizan un software estadístico desde casi el primer día de clase y muchas veces resuelven problemas del ámbito farmacológico con el software sin necesidad de conocer en profundidad la técnica estadística apropiada. Creemos que, si en el grado de Farmacia, así como en otros grados, la experiencia ha sido positiva, en los grados en informática con mayor motivo ya que los alumnos buscan un grado aplicado y conocer distintos softwares que será lo que usen en el futuro cuando se enfrenten al mercado laboral.

Todavía no hemos podido valorar los resultados ya que los alumnos todavía no se han examinado. Sin embargo, sí podemos decir que el porcentaje de alumnos que han participado en la actividad que este año hemos ofrecido como optativa en el grado de Ingeniería del Software ha sido de entre el 40 y 50% del total de matriculados. Estos alumnos han elegido ellos mismos la base de datos a tratar y han resuelto el análisis descriptivo y de regresión con éxito.

3. Metodología empleada en el proyecto

Los miembros del proyecto pertenecen al departamento de Estadística e Investigación Operativa de la Facultad de Matemáticas. Las asignaturas implicadas en el proyecto pertenecen a los distintos grados de la Facultad de Informática. El grupo de profesores que solicitamos este proyecto hemos venido notando que la integración adecuada de un software estadístico aplicado favorece una mayor participación, creatividad, aprendizaje y motivación por parte del estudiante. El uso de este software permitiría al estudiante comprender las ideas, conceptos y aplicaciones de la asignatura.

Los profesores que nos hemos planteado este proyecto llevamos algún tiempo impartiendo las distintas asignaturas de *Estadística* para los graduados en las distintas titulaciones de Ingenieros Informáticos. Por ello, en primer lugar hemos tenido varias reuniones de coordinación, lo que nos llevó a la petición de este proyecto.

De forma más específica, los cuadernillos elaborados tendrán una parte resumen de cada tema o bloque visto en clase y una explicación del enfoque computacional. Esta segunda parte consistirá en un código o un conjunto de instrucciones para ejecutar con el ordenador la técnica vista en clase. Además, habrá una serie de ejercicios relacionados para poner en práctica. De esta forma, el tiempo en el ordenador servirá siempre de refuerzo a la asignatura.

La metodología a seguir consiste en la elaboración de guías de estudio o cuadernillos de prácticas. Estos cuadernillos tendrían una doble función, por una parte permitir el autoaprendizaje del alumno, teniendo siempre al profesor para consulta y guía y, por otra, servir para las pruebas de evaluación continua.

La idea inicial es que el profesor dedique dos sesiones de una hora a explicar el cuadernillo (repaso de cada tema o bloque) en el aula. No es preciso hacerlo en el laboratorio, basta con que el profesor abra el programa de trabajo y el cuadernillo para ir explicando la técnica correspondiente con su ordenador. Después será el alumno el que por su cuenta (bien en grupos o bien de forma individual) resuelva parte o todo el cuadernillo. La resolución de los ejercicios del cuadernillo por parte del alumno será lo que constituya parte de la evaluación continua.

4. Recursos humanos

Los profesores involucrados en el proyecto son 5. Todos los integrantes cuentan con evaluaciones positivas de docencia. En particular, la investigadora principal ha obtenido una evaluación trienal EXCELENTE en la convocatoria 2016-2017 y un premio de excelencia. Además, todos los miembros han sido profesores de las distintas asignaturas de Estadística en la Facultad de Informática. Concretamente, el profesor Vadillo cuenta con la experiencia de haber coordinado la asignatura “Probabilidad y Estadística” en la titulación de Ingeniería Informática (con un grupo en inglés) y este año es coordinador de la asignatura en los grados de Ingeniería Informática, Ingeniería de los Computadores y el reciente grado de Desarrollo de Videojuegos. La profesora Ramos también ha coordinado durante varios cursos la asignatura “Estadística Aplicada” en el grado de Ingeniería del Software.

También contamos en el grupo investigador con la colaboración de la profesora Landáburu, que lleva impartiendo esta asignatura desde el comienzo de los grados adaptados al EEES. La aportación de sus experiencias en los distintos cursos ha sido de mucha utilidad a la hora de decidir la solicitud de este proyecto. Otros miembros del proyecto son el profesor Juan José Prieto y Oscar Vadillo. Profesores que también imparten docencia de Estadística en los grados de Informática.

Creemos que cada uno puede aportar su visión y experiencia y contribuir, así, de forma eficiente a la elaboración del material. Siempre con la visión puesta en los objetivos marcados, principalmente el poder llegar a más estudiantes, haciendo de la asignatura una asignatura más atractiva y mejorando así la tasa de éxito.

El mismo grupo de profesores hemos solicitado un nuevo proyecto para continuar elaborando el banco de ejercicios, usando esta vez el tema de Probabilidad. Ya que el proyecto también fue concedido esperamos tener elaborado hacia finales del próximo curso un manual de uso y banco de ejercicios considerable. Una vez concluido dicho manual, se solicitará la edición como libro electrónico a través de los e-prints de la Universidad Complutense.

5. Desarrollo de las actividades

Para la puesta en marcha del proyecto se plantearon las siguientes actividades:

- 1.- Reunión del equipo del proyecto, donde se abordaron los distintos tipos de software a utilizar y los distintos problemas a tratar desde un punto de vista descriptivo.
- 2.- Elaboración de un manual que recoja los procedimientos estadísticos a tratar en los problemas desde un punto de vista descriptivo.
- 3.- Elaboración de hojas de ejercicios de descriptiva y regresión, conjuntos de datos, etc.
- 4.- Puesta en práctica con el ordenador en el aula del análisis de bancos de datos con el software elegido.
- 5.- Reunión del equipo para proponer el trabajo optativo que constituya parte de la nota de evaluación continua.
- 6.- Valoración de los trabajos presentados y evaluación de los mismos.

6. Anexos

Presentamos en este anexo una parte del cuadernillo utilizado por los alumnos para la elaboración de los problemas que debían tratar y entregar resueltos. El cuadernillo en su totalidad, tal y como ha sido facilitado a los alumnos, está colgado en el campus virtual y a disposición del Vicerrectorado de Calidad si lo considera oportuno.

También mostramos a modo de ejemplo algún ejercicio resuelto que les sirviera de guía para la resolución de los problemas que ellos abordaban.

Se evaluó positivamente la búsqueda de datos relacionados con los temas de la Ingeniería Informática así como la resolución a través del Excel del análisis descriptivo de los datos elegidos. Los mejores ejercicios realizados servirán como ejemplos para las nuevas generaciones de las titulaciones en Informática.

Cuaderno de prácticas de Estadística Descriptiva en Excel

ÍNDICE

1. Introducción a Excel

- 1.1 Ficheros de Excel
- 1.2 Celdas
- 1.3 Gráficos

2. Estadística descriptiva

- 2.1 Variable discreta. Tabla de frecuencias
- 2.2 Variable discreta. Gráficos
- 2.3 Variable discreta. Principales estadísticos
- 2.4 Variable continua. Tabla de frecuencias
- 2.5 Variable continua. Gráficos
- 2.6 Variable continua. Principales estadísticos

3. Modelo de regresión

1. INTRODUCCIÓN A EXCEL

Excel es un potente programa de hojas de cálculo integrado en el conjunto ofimático de programas Microsoft Office que permite, entre otras cosas:

- Organizar, modificar y analizar datos
- Realizar cálculos rápidos y complejos
- Tomar decisiones a partir de la ejecución de complementos
- Obtener Gráficos diversos a partir de datos

Una hoja de Cálculo no es una aplicación específicamente estadística. Sin embargo, contiene herramientas de tipo estadístico, de acceso generalizado, que son suficientes para realizar estadísticas básicas.

Aunque los paquetes estadísticos presentan características más ventajosas, nos permiten ejecutar procedimientos estadísticos más avanzados, los vamos a sustituir por la facilidad de acceso del Excel y por su aplicabilidad y posibilidad de generalización a diversos campos de trabajo, no sólo la Estadística sino la Optimización, la Contabilidad,...

1.1 FICHEROS EXCEL

Los ficheros de Excel se denominan libros. Cada libro está compuesto por diferentes hojas de cálculo (hasta 255). Por defecto se crea con tres hojas.

Cada hoja de cálculo está formada por una matriz de celdas que almacenan información. Esta información puede ser de tipo texto (conjunto de caracteres), numérico (números propiamente dichos, porcentajes y fechas) o fórmulas (expresiones que comienzan con el símbolo =, y donde se pueden mezclar constantes, nombres de celdas, referencias a otras celdas, operadores y funciones).

Cada hoja tiene hasta 16384 columnas y 1.048.576 filas. Las filas se identifican de forma consecutiva por números 1, 2, 3,... y las columnas tienen asociadas letras simples A, B, C,... dobles AA, AB,... o triples. Cada celda se nombra por una fila y columna. Por ejemplo, B4 se corresponde con la celda que está en la cuarta fila y la segunda columna.

Para arrancar Excel debemos seleccionar:

Inicio ⇒ Programas ⇒ Microsoft Office ⇒ Microsoft Excel

1.2 CELDAS

Edición de una celda

En las celdas se hacen operaciones, transformaciones o cálculos de diversos tipos. Para cambiar el contenido de una celda debes seleccionarla previamente (convertirla en celda activa). Puedes escribir directamente en la celda y si ésta tiene ya contenido, lo sustituye. También puedes escribir en la barra de fórmulas y pulsar Enter. Si pulsas sobre la tecla con doble clic, permite corregir.

Si pulsas Ctrl mientras seleccionas varias celdas. Escribes un contenido y pulsas Ctrl+Intro se rellenarán con el mismo valor. Si editas una celda y pinchas en la esquina derecha de la misma y arrastras a las demás, se rellenarán con el mismo valor. Si es un mes, aparecerán los restantes meses del año. Si quieres introducir una sucesión como: a1, a2,... ó 1, 2,... ó 1, 3, 5,... deberás seleccionar las dos primeras celdas y pinchar en la esquina derecha y arrastrar a las siguientes celdas.

Borrado de una celda

Se debe distinguir entre:

- *Borrar celdas:* Se eliminan totalmente las celdas y otras ocupan su lugar. El programa solicita información sobre cuáles serán las celdas que reemplazarán a las eliminadas. Esta operación se ejecuta mediante el botón Eliminar de la ficha de Inicio. Te da opción a elegir entre eliminar celdas, filas o columnas completas. Hay que tener cuidado al elegir las opciones posteriores de cómo desplazar las celdas vecinas.
- *Borrar contenido:* En el segundo caso, se debe concretar si se borra todo, o sólo formato, hipervínculos, comentarios, etc. Esta operación se puede hacer con la tecla Supr (para borrar contenidos) o con el botón Borrar situado a la derecha del botón Eliminar y después se especifica si se quiere borrar todo o algún elemento.

Referencias a una Celda

Toda celda tiene un nombre, una referencia, que la distingue de las demás.

Referencia Relativa. En operaciones de copiar o rellenar rangos el Excel supone que al mover las celdas también mueve su contenido. Los datos sufren el mismo movimiento que la celda que los contiene.

Referencia Absoluta. Al mover la celda no altera ningún dato que contenga. Van precedidas por el signo \$. Hay tres modalidades:

- Tipo \$B\$30: es totalmente absoluta. No se altera ni la fila ni la columna.
- Tipo \$B30: sólo se protege la columna, pero se puede alterar la fila.
- Tipo B\$30 se alterará la columna y la fila permanecerá sin cambios.

Para convertir una referencia de relativa a absoluta se pulsa F4.

Formato de una Celda:

El formato de una celda o rango de celdas es el conjunto de opciones que constituyen su forma de presentación en pantalla (el conjunto de tipos de letras, colores, alineación o bordes que cambian el aspecto de la celda) y que no afectan a su contenido.

A todas esas opciones se accede al abrir la ficha Inicio o pulsando el botón derecho del ratón en el menú contextual Formato de celdas.

El formato de una celda puede ser de uno de los siguientes tipos:

1. Número: Las opciones de formato numérico son muy variadas.
2. Alineación: Permite cambiar el contenido de una celda respecto a su contorno.
3. Fuente: Sirve para seleccionar el tipo de letra, el estilo (negrita, cursiva), el tamaño, el color, los efectos (subíndice, superíndice, tachado), etc.
4. Bordes: Permite cambiar los bordes de una celda o grupo de celdas. Se pueden cambiar el estilo de línea, los colores, los bordes predeterminados.
5. Relleno: Sirve para cambiar el color del fondo de una celda, color de la trama, estilo de la trama, etc.
6. Proteger: Permite proteger las celdas de una hoja para evitar cambiar su contenido. La protección sólo se activará si en la Ficha Revisar en el bloque Cambios se ha elegido Proteger hoja.

Ordenar y Filtrar un rango de Celdas:

Para ordenar un rango de celdas previamente se ha de seleccionar y luego pulsando en el botón Ordenar y Filtrar se puede elegir entre:

- Alfabético de la A a la Z
- Alfabético de la Z a la A
- Orden Personalizado

Para filtrar o seleccionar determinadas celdas que verifican un criterio, primero se seleccionan y se pulsa en el botón Ordenar y Filtrar y luego en filtro. Entonces se añade una flechita sobre las celdas y pulsando sobre ella, se abre un menú contextual. Si el contenido de las celdas es numérico se habla de Filtro de Números para elegir el criterio y si es de texto de Filtro de texto.

Las fórmulas se escriben de forma algebraica, usando los operadores +, -, *, / y ^. Se deben usar paréntesis para indicar la prioridad de operaciones. Estos operadores se aplican a:

- Números: =30*2^3
- Referencias y nombres de celdas: =C5/3+C4/2; =TOTAL*(23+12*C4)
- Funciones: =SUMA(E3:E20)+D10

También se admiten operadores de comparación: <, >, <=,

Toda fórmula ha de venir precedida del signo = para que Excel interprete que es una expresión que debe calcular. Hay que distinguir entre la fórmula y su resultado, que es lo que aparece en

pantalla. Así, si escribimos por ejemplo 22*2+2, el resultado será 46 que es lo que mostrará en pantalla.

Las funciones son una herramienta de uso común. Para activar una función vamos a:

Fórmulas \Rightarrow Insertar Función

o pulsamos el botón f_x situado en la barra de fórmulas. Las fórmulas están compuestas de unas Palabras Clave seguida de un Argumento (datos sobre las que se aplican) escrito entre paréntesis. Las más elementales son: SUMA, CONTAR, PROMEDIO, SI, RAIZ, ALEATORIO,...

Ejemplos:

- =SUMA(C4:D12) Esta función sumaría todos los números comprendidos entre la celda C4 y la D12. Aquí SUMA es la palabra clave y C4:D12 el argumento.
- =HOY() Devuelve la fecha actual. No tiene argumentos, pero se deben escribir los paréntesis.
- =RAIZ(9) Devolvería un 3, que es la raíz cuadrada de 9. RAIZ palabra clave y 9 el argumento.
- =PROMEDIO(C4:G18) calcula la media aritmética de todos los números contenidos en las celdas comprendidas entre C4 y G18.

1.3 GRÁFICOS

Excel permite representar una variedad de gráficos a partir de una serie de datos.

Columna	Área	Anillos
Línea	Dispersión	Burbuja
Circular	Cotizaciones	Radial
Barra	Superficie	

Para insertarlos se selecciona:

Insertar \Rightarrow Gráfico

Primero seleccionamos los datos, o una celda vacía, y elegimos en Herramientas de Gráficos en la pestaña de Diseño, Seleccionar datos y eligiendo el rango de celdas a representar con el ratón.

Ejercicio.

La corrosión del acero de refuerzo es un problema grave en las estructuras de cemento situadas en lugares expuestos a condiciones meteorológicas severas. Por esta razón, un grupo de ingenieros están investigando el uso de barras de refuerzo hechas con un material compuesto con fibra de vidrio. Para analizar las ventajas de este producto, se recogieron las 48 observaciones sobre la resistencia de la mezcla que figuran en la hoja corrosión. a) Resume estos datos en una tabla de frecuencias. b) Realiza un gráfico adecuado para estos datos. c) Calcula la media muestral y la mediana muestral de estos datos. ¿Qué medida resumen resulta más adecuada en este caso? ¿Por qué? d) Calcula los cuartiles muestrales de los datos anteriores. Interpreta los resultados.

Solución al ejercicio propuesto.

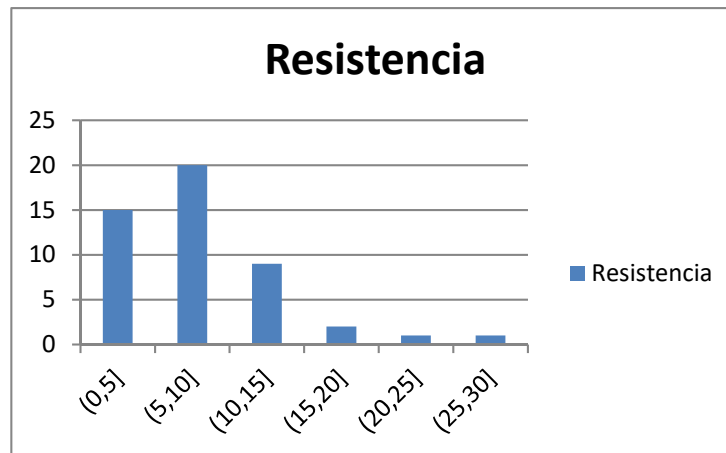
X= "resistencia de la mezcla" Rango: 3,4 25,5

	Valores	fr. Abs ni	fr. Rel fi	fr. Abs. Acum Ni	fr. Rel. Acum Fi
11,5	(0,5]	15	0,3125	15	0,3125
12,1	(5,10]	20	0,41666667	35	0,729166667
9,9	(10,15]	9	0,1875	44	0,916666667
9,3	(15,20]	2	0,04166667	46	0,958333333
7,8	(20,25]	1	0,02083333	48	0,979166667
6,2	(25,30]	1	0,02083333	1	1
6,6		48	1		

- a) Tengo que copiar los límites de los intervalos.
Ojo porque el último extremo y el primero no los pongo.
Luego selecciono toda la columna de "fr.",
uso comando FRECUENCIA(datos; grupos) y marco CONTROL+SHIFT+INTRO
- 5,7
- 5,4
- 5,2
- 5,1
- 4,9
- 10,7
- 15,2
- 8,5
- 4,2
- 4
- 3,9
- 3,8
- 3,6
- 3,4
- 20,6
- IMPORTANTE: Para las fórmulas matriciales,
1) Escribo la fórmula y doy a ENTER (la solución aparece en una celda)
2) Selecciono las casillas que van a formar la matriz de resultados
3) Presiono F2 (de esa forma aparece la fórmula entre {})
4) Presionar CTRL+SHIFT+ENTER.

- b) Gráfico columna = histograma

25,5
13,8
12,6
13,1
8,9
8,2
10,7
14,2
7,6
5,2
5,5
5,1
5
5,2
4,8
4,1
3,8
3,7
3,6
3,6
3,6



c) Media Muestral de X: 8,07916667
Mediana Muestral de : 5,95

La medida más adecuada sería la media. Hay cuatro datos con valores grandes alejados del resto que hacen que la media resulte alta.

Sin embargo, la mayor parte de los datos tienen una resistencia entre 5 y 10.

d) Primer Cuartil Muestral de X	4,65	PERCENTIL	k en (0,1)
Segundo Cuartil Muestral de X (mediana)	5,95		
Tercer Cuartil Muestral de X	10,7		

El primer y segundo cuartil muy próximos porque es donde se concentran más los datos.